Received: 2 May 2020 | Revised: 11 June 2020 | Accepted: 13 June 2020

DOI: 10.1002/dvg.23386

RESEARCH ARTICLE

genesis WILEY

Diphtheria toxin-mediated transposon-driven poly (A)-trapping efficiently disrupts transcriptionally silent genes in embryonic stem cells

Jie Bai[1,2] | Ryohei Kondo[1,3] | N. Ika Mayasari[1,4] | Toshiaki Shigeoka[1,5] | Ayako Isotani[6,7] | Masahito Ikawa[8] | Goro Sashida[2] | Masashi Kawaichi[9] | Yasumasa Ishida[1]

[1]Laboratory of Functional Genomics and Medicine, Division of Biological Science, Nara Institute of Science and Technology, Nara, Japan

[2]Laboratory of Transcriptional Regulation in Leukemogenesis, International Research Center for Medical Sciences, Kumamoto University, Kumamoto, Japan

[3]Department of Peripheral Nervous System Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan

[4]Faculty of Veterinary Medicine, Bogor Agricultural University (IPB), Jalan Agatis Kampus IPB Darmaga, Bogor, West Java, Indonesia

[5]Wellcome Trust-Medical Research Council Stem Cell Institute, University of Cambridge, Cambridge, United Kingdom

[6]Immunology Frontier Research Center, Osaka University, Osaka, Japan

[7]Organ Developmental Engineering, Division of Biological Science, Graduate School of Science and Technology, Nara Institute of Science and Technology, Nara, Japan

[8]Research Institute for Microbial Diseases, Osaka University, Osaka, Japan

[9]Division of Educational Development, Nara Institute of Science and Technology, Nara, Japan

Correspondence
Jie Bai and Yasumasa Ishida, Laboratory of Functional Genomics and Medicine, Division of Biological Science, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma-shi, Nara 630-0192, Japan.
Email: j-bai@kumamoto-u.ac.jp (J. B.) and ishiday@bs.naist.jp (Y. I.)

Funding information
Japan Society for the Promotion of Science, Grant/Award Numbers: 19310130, 21310128; Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan, Grant/Award Numbers: H19-20, H21-22

Summary

Random gene trapping is the application of insertional mutagenesis techniques that are conventionally used to inactivate protein-coding genes in mouse embryonic stem (ES) cells. Transcriptionally silent genes are not effectively targeted by conventional random gene trapping techniques, thus we herein developed an unbiased poly (A) trap (UPATrap) method using a *Tol2* transposon, which preferentially integrated into active genes rather than silent genes in ES cells. To achieve efficient trapping at transcriptionally silent genes using random insertional mutagenesis in ES cells, we generated a new diphtheria toxin (DT)-mediated trapping vector, DTrap that removed cells, through the expression of DT that was induced by the promoter activity of the trapped genes, and selected trapped clones using the neomycin-resistance gene of the vector. We found that a double-DT, the dDT vector, dominantly induced the disruption of silent genes, but not active genes, and showed more stable integration in ES cells than the UPATrap vector. The dDT vector disrupted differentiated cell lineage genes, which were silent in ES cells, and labeled trapped clone cells by the expression of EGFP upon differentiation. Thus, the dDT vector provides a systematic approach to disrupt silent genes and examine the cellular functions of trapped genes in the differentiation of target cells and development.

KEYWORDS
cell lineage ablation, diphtheria toxin, embryonic stem cells, gene trap, *Tol2* transposon

1 | INTRODUCTION

Mutagenesis in cells or animals is one of the genetic methodologies employed to elucidate the molecular functions of genes regulating biological processes. Random gene trapping is the application of insertional mutagenesis techniques that are conventionally used to inactivate protein-coding genes in mouse embryonic stem (ES) cells (Stanford, Cohn, & Cordes, 2001). Gene trapping was utilized in the knockout mouse project (KOMP) because it was easier and cheaper than classical gene knockout when the number of target genes was large (Austin et al., 2004). The aim of random gene trapping is to disrupt target genes by two major methods, promoter trapping and poly (A) trapping. While a promoter-trapping vector harboring a promoter-less selection cassette generally disrupts transcriptionally active genes, but does not typically capture silent genes in target cells (Gossler, Joyner, Rossant, & Skarnes, 1989), a poly (A)-trapping vector drives and stabilizes selection-cassette mRNA by adding a poly (A) signal driven from the region of the trapped gene, regardless of the transcription status of the trapped gene in cells (Ishida & Leder, 1999; Niwa et al., 1993). We generated an original poly (A)-trapping retrovirus vector of UPATrap, which suppressed the inappropriate activation of the nonsense-mediated mRNA decay (NMD) pathway and targeted both transcriptionally active and silent genes in the unbiased integration of the virus vector in genomes (Shigeoka, Kawaichi, & Ishida, 2005). To improve the stability and integrity of the integrated vector in cells, we utilized a cut and paste-type DNA transposon, *Tol2* in the backbone of the vector (Kawakami, Shima, & Kawakami, 2000; Koga, Suzuki, Inagaki, Bessho, & Hori, 1996; Urasaki, Morvan, & Kawakami, 2006), and successfully achieved the conditional disruption of the gene in cells via this unbiased poly (A)-trapping vector (Mayasari et al., 2012).

Moloney murine leukemia virus (MMLV) vectors were previously shown to preferentially integrate into transcriptionally active genes (Scherdin, Rhodes, & Breindl, 1990; Wu, Li, Crise, & Burgess, 2003), and we also demonstrated that the UPATrap retrovirus vector frequently integrated into active genes, but only approximately 10% of all trapped genes, which were transcriptionally silent in murine ES cells (Mayasari et al., 2012). By using the *Tol2*-driven UPATrap vector, we achieved a higher trapping frequency in silent genes of approximately 25% of all trapped genes, which was markedly smaller than the expected trapping frequency of 45% in silent genes in ES cells (Mayasari et al., 2012), indicating the preferential integration of the vector in active genes.

To efficiently target silent genes in cells and examine the cellular functions of disrupted genes upon differentiation using the *Tol2*-driven UPATrap vector, we herein adopted a negative selection strategy, by which diphtheria toxin (DT) produced in *Corynebacterium diphtheria* conferred strong cytotoxicity by inactivating the eukaryotic polypeptide elongation factor, resulting in the inhibition of protein synthesis (Kohno et al., 1986; Kohno & Uchida, 1987). We generated a new DT-mediated *Tol2*-driven trapping vector, DTrap, and found that a derivative DTrap vector of double DT containing the dDT vector dominantly induced the disruption of silent genes, but not active genes, and showed more stable integration in ES cells than the original UPATrap vector. We also demonstrated that the dDT vector disrupted differentiated cell lineage genes, which were silent in ES cells, and labeled the trapped clone by the expression of EGFP upon differentiation.

2 | MATERIALS AND METHODS

2.1 | Gene trapping vectors

A wild-type DT-A cassette was cloned into the ClaI-BamHI site of a *Tol2* transposon CTP2F vector to create a DTrap-CTP2F vector. A gene-terminator cassette containing a promoter-less EGFP was inverted into the DTrap-CTP2F vector to generate a sDT vector containing a single DT cassette. An inverted second DT cassette was inserted downstream of the NEO cassette to create a double DT-containing dDT vector. A weaker toxin DT176 vector, the point mutation of which substituted Gly to Asp at residue 128 of the DT protein, was cloned by replacing the 0.3-kbp BstZ17I-BlpI fragment of the DT vector with the corresponding fragment of the pCRM176 vector (Uchida, Pappenheimer Jr., & Greany, 1973). These DTrap vectors contained inverted pairs of the *FRT* and *F3* sequences for FLPo-mediated recombination and an identification cassette containing either of fifteen different tag sequences (#01 to #15). The sequences of the sDT, dDT, sDT176, and dDT176 vectors harboring the #01 tag were deposited under the GenBank/EMBL/DDBJ accession numbers LC085658, LC085659, LC085660, and LC085661.

2.2 | Vector plasmid transfection

A total of 2.5×10^5 ES cells were transfected with 1.25 μg of pCAGGS-TP, which codes the *Tol2* transposase (Kawakami et al., 2000), and 0.125 μg of each mixture of differentially-tagged sDT, dDT, sDT176, or dDT176 using the TransFast reagent (Promega). After the treatment with 200 μg/ml of G418 (Nacalai) for 10–13 days (Matsuda et al., 2004; Shigeoka et al., 2005), neomycin-resistant colonies were isolated and expanded in an in vitro culture. Genomic DNA and total RNA were extracted from the expanded clones.

2.3 | ES cell culture and generation of chimeric mice

V6.4 and KY1.1 ES cells were cultured on mitomycin C-treated SNL-STO cells, which were infected using a retrovirus vector stably

expressing murine leukemia inhibitory factor (LIF) and the neomycin-resistance gene product (McMahon & Bradley, 1990). A blastocyst microinjection was used to produce chimeric mice. All experiments using these mice were performed in accordance with institutional guidelines for the use of laboratory animals and approved by the Review Board for Animal Experiments of Nara Institute of Science and Technology University.

2.4 | In vitro differentiation of ES cells

To induce the differentiation of adipocytes, ES cells were cultivated in aggregates termed embryo bodies (EBs) and hanging drops containing 1,000 ES cells in 20 μl of cultivation medium were maintained for 2 days on the lids of bacteriological dishes. EBs were maintained for 3 days in medium supplemented with 1 μM all-trans retinoic acid (Sigma) and then treated in differentiation medium with 100 nM insulin (Sigma), 2 nM triiodothyronine (Sigma), 0.1 mM 3-Iaobutyl-1-methylxanthine (Sigma), and 10 nM dexamethasone (Wako) (Dani et al., 1997; Rubin, Hirsch, Fung, & Rosen, 1978).

2.5 | FLPo-mediated recombination in ES cells

ES cell clones were transfected with a pCAGGS-FLPo-IRES-Puro resistant-poly (A) vector, and treated with 1 μg/ml puromycin for 48 hr for selection. Single colonies were isolated using a limiting dilution method and expanded for a 6–8-day culture on a layer of mitomycin C-treated SNL-STO cells. A total of 24–36 FLPo-generated daughter sub-clones were selected to confirm the efficacy of FLPo-mediated recombination by genomic PCR. We amplified between the intra-vector and intragenic regions using the primer sequences shown in Table S1.

2.6 | Assessment of the number and direction of vector integration and integrated vector integrity

Genomic DNA was extracted from ES cell clones and amplified by PCR using Phusion Hot Start High-Fidelity DNA polymerase (Thermo Scientific). Primers for PCR and sequencing were shown in Table S1. Genome-integrated vector integrity was examined by PCR using KOD-FX (TOYOBO) polymerase. Five regions of the vectors were amplified using five distinct primer sets (Table S1).

2.7 | Splinkerette genome PCR

To identify the integration sites of the transposon, genomic DNA isolated from ES cell clones was digested with HaeIII, TaqI, or MspI (New England BioLabs). After the inactivation of restriction enzymes, digested DNA was ligated with compatible splinkerette-type linkers for HaeIII-digested DNA or TaqI- and MspI-digested DNA using T4 DNA ligase (Takara). Two rounds of PCR using KOD-FX polymerase (TOYOBO) and distinct primer sets were performed on ligated DNA and direct sequencing was then conducted to detect the vector-integrated region. The sequences of the linkers and primers are shown in Table S1.

2.8 | Gene expression dataset analysis

The expression levels of genes in ES cells were assessed using NCBI dbEST libraries (#1882, #2512, #10023, #14556, #15703, #17907, and #21037) (Mayasari et al., 2012) and RNA-sequencing data in V6.5 ES cells (GSM521650) (Guttman et al., 2010). Sequence reads were mapped to the mouse Refseq gene using the Burrows-Wheeler Alignment tool (BWA) with default settings (Li & Durbin, 2009) and Reads Per Kilobase of exon per million (RPKM) values were calculated using Artemis (Carver, Harris, Berriman, Parkhill, & McQuillan, 2012).

2.9 | Quantitative RT-PCR

Total RNA was isolated using the RNeasy Plus Mini kit (Qiagen) and cDNA was generated by oligo dT or random hexamers using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen). Quantitative RT-PCR (qRT-PCR) was performed using gene-specific primers and the Thunderbird qPCR Mix (TOYOBO) in a LightCycler 96 System (Roche). Primer sequences are shown in Table S1.

2.10 | Oil red O staining

Cells were washed with PBS and fixed with 10% paraformaldehyde at room temperature for 10 min, and then washed twice with PBS followed by washing with 60% isopropyl-alcohol for 1 min. Fixed cells were stained with 60% Oil Red O solution from a stock of 150 mg Oil Red O (Sigma) in 50 ml isopropyl-alcohol for 20 min.

3 | RESULTS

3.1 | Generation of DTrap vectors targeting transcriptionally silent genes

To disrupt genes that are transcriptionally silent in ES cells, we initially generated new DTrap vectors designated for poly (A) trapping and subsequent negative selection (e.g., sDT, dDT, sDT176, and dDT176 vectors) (Figure 1a,b). In these vectors, the DT cassette consists of an efficient splice-acceptor sequence (Ishida & Leder, 1999), and an IRES-driven DT coding region for wild-type DT (sDT) or attenuated DT176 (sDT176), followed by poly (A)-addition signals (Figure 1a). We also generated the double DT cassette-containing vectors, dDT and dDT176, which disrupt target genes regardless of the trapping orientation of vectors (Figure 1b). The attenuated DT-A gene (*tox* 176) contains a G-to-A transition at nucleotide 383 that results in the

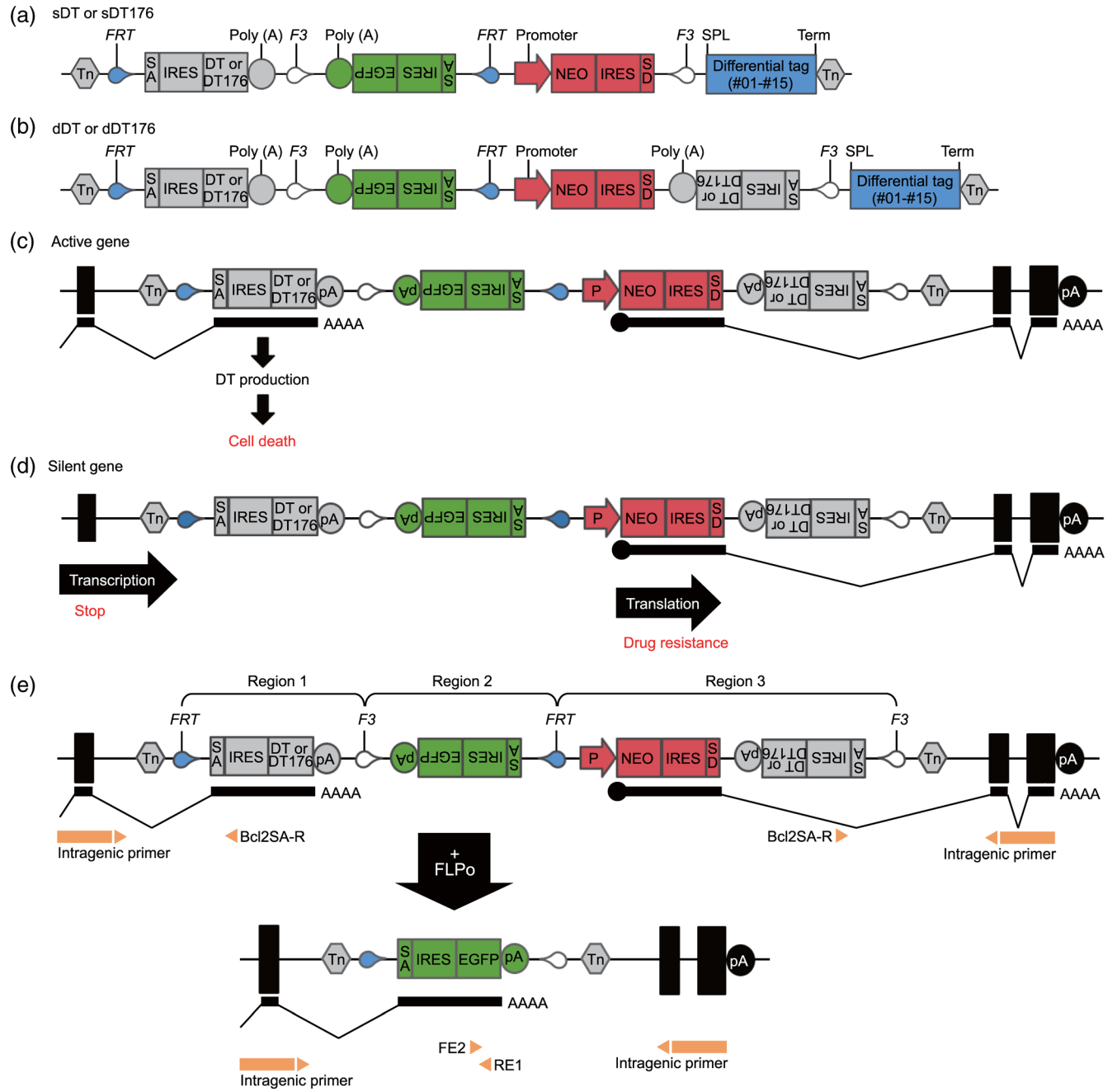

FIGURE 1 Generation of the DT-mediated UPATrap vector. (a) A single DT (sDT) vector harboring wild-type DT or an attenuated tox-176 (DT176) cassette (gray box) with 15 differential tag sequences (blue box). (b) A double DT (dDT) vector harboring two wild-type DT or two attenuated tox-176 (DT176) cassettes (gray boxes) with 15 differential tag sequences (blue box). (c) DT vectors containing the diphtheria toxin that induced the death of cells integrated at an active gene in the forward orientation. (d) DT vectors selecting cells in which the vector is integrated at a silent gene by lacking the induction of toxins, but expressing the neomycin-resistant gene (red box). (e) Induction of FLPo-mediated homologous recombination at the DT vector-integrated region through the Flp recombinase-target signals of *FRT* (blue teardrops) and *F3* (opened teardrops). Orange triangles and arrows show the PCR primers used to confirm recombination. (a–e) Tn, SA, SD, and SPL stand for the terminal essential sequences (L200 and R175) of the *Tol2* transposon, splice-acceptor sequence of the intron 2-modified exon 3 of the human *BCL-2* gene, splice donor sequence of the modified exon 8-intron 8 of the murine *Hprt* gene, and synthetic nucleotide sequence amplified by splinkerette genome PCR, respectively. Black closed boxes and lines represent the exons of a trapped gene and the exon and intron portions of pre-mRNA

replacement of glycine at position 128 by aspartic acid (Maxwell, Maxwell, & Glode, 1987). The enzymatic activity of attenuated DT-A was 30- to 100-fold smaller than that of wild-type DT-A, as assessed in human 293 cells (Maxwell et al., 1987; Yamaizumi, Mekada, Uchida, & Okada, 1978). The second portion of the gene-terminator cassette consists of an efficient splice acceptor sequence, IRES-driven enhanced green fluorescent protein (EGFP) cDNA, and four copies of the poly (A)-addition signals. To completely terminate pre-mRNA

transcription, the gene-terminator cassette was constructed in an inverted configuration in these vectors (Figure 1a,b). The third portion is a constitutively active promotor-driven Neomycin-resistant gene (NEO) followed by a poly (A)-trapping cassette, which abrogates the activation of nonsense-mediated mRNA decay (Figure 1c) (Mayasari et al., 2012; Shigeoka et al., 2005). When these vectors were integrated inside an active gene in the correct orientation, the endogenous promoter activity of the trapped gene drove the expression of DT (or DT176), which resulted in the death of the targeted cell (Figure 1c). In contrast, the vectors integrated in a transcriptionally silent gene did not initiate the expression of DT (or DT176), but activated the expression of NEO, which led to the selection of silent gene-trapping cells (Figure 1d). Furthermore, single 5′-oriented DT cassette vectors, such as sDT and sDT176, were integrated in an active gene in the inverted orientation, resulting in the failed selection of targeted cells, whereas double DT cassette-vectors selected the targeted cells via the trapping gene regardless of the orientation of the vector upon integration (Figure S1).

To trace the disruption of genes in cells after the selection of individual clones, we introduced the target signals of Flp recombinase (*FRT* and *F3*) at four sites in the vectors (Figure 1a,b). When the transient expression of Flp recombinase induced FlEx-type recombination within a genome-integrated vector region in the cell, the 5′ DT cassette, Neo-poly (A)-trapping cassette, and second DT cassette were removed from the genome in the cell, resulting in the inverted recombination of the gene-terminator cassette (Figure 1e). Thus, we successfully generated new DTrap vectors, which were designated to efficiently trap silent genes and select trapped clones.

3.2 | Weaker toxin DT176 vectors efficiently trapped at a single gene-coding region in ES cells

To assess the trapping efficacy and integrating accuracy of DTrap vectors, we initially transduced a mixture of differentially-tagged vectors, which harbor 15 unique sequences in the last 3′ region (Figure 1a,b), into ES cells. We did not observe any ES cell clone containing more than three transposons in PCR and direct sequencing experiments, while the dDT176 vector showed larger single-trapping clones than the sDT176 vector (80.8% versus 67.4%) (Figure 2a). Among the 15 differentially tagged vectors in sDT176- and dDT176-transfected ES cells, we found a weak bias in a few subsets of the sDT176 vector, whereas all of the other vectors were similarly transduced in ES cells after the selection of clones (Figure 2b). To evaluate the efficacy of the selection of clones of DTrap vectors, we performed gene trapping

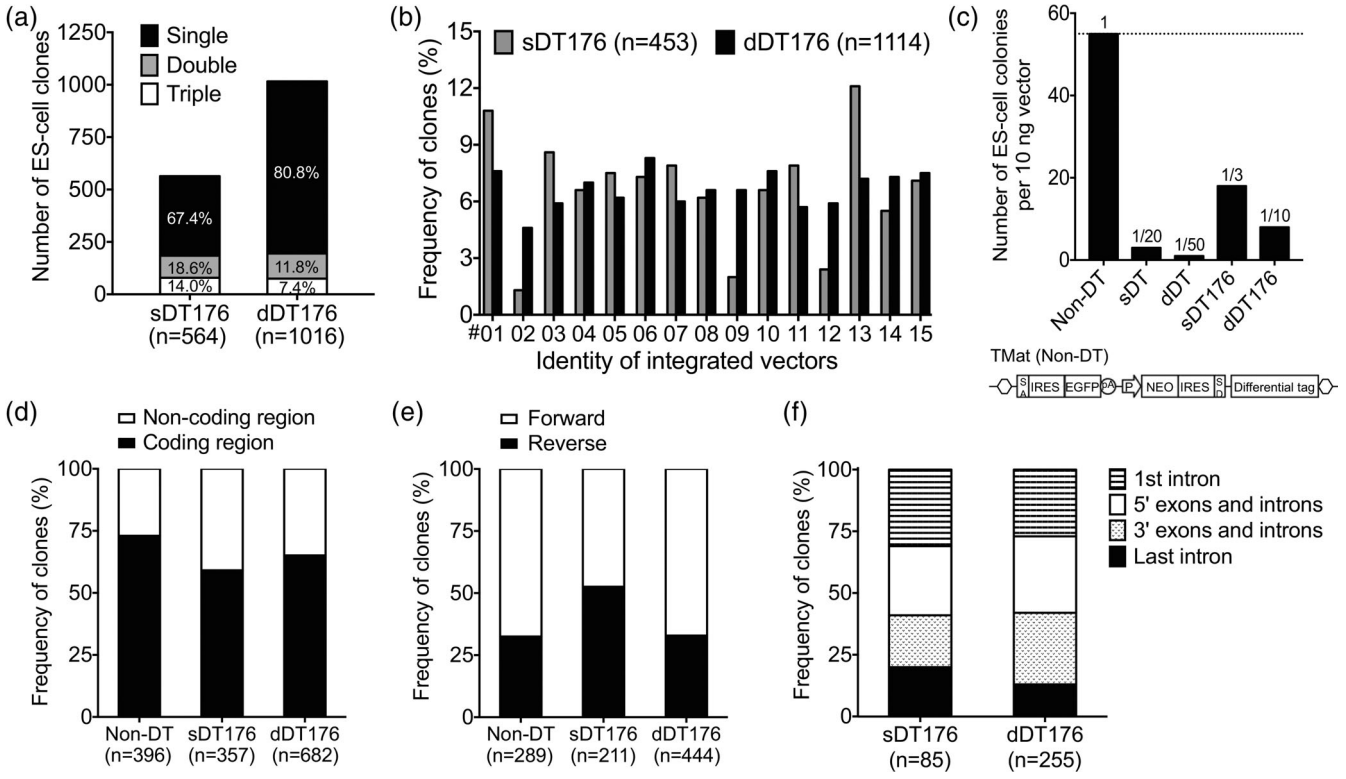

FIGURE 2 Weaker toxin DT176 vectors efficiently trapped at a single region in stem cells. (a) Number of ES cell clones harboring one to three integration site(s) of the DT176 vector. (b) Similar integration frequency of DT176 vectors among 15 differentially-tagged clones. (c) Numbers of ES cell colonies obtained by transfecting an equal amount of 10 ng DNA of the vector plasmid and a non-DT TMat vector, the structure of which was shown at the bottom. (d,e) Frequencies of the integration regions of (d) and integration orientations of (e) the vectors assessed by performing splinkerette genome PCR at the SPL region shown in Figure 1a,b. (f) Distribution of DT176 vector integration regions between the first intron, 5′ half exons and introns, 3′ half exons and introns, and the last intron in trapped genes showing a forward-integration orientation

in ES cells using the DTrap vector and UPATrap vector, which does not contain the DT domain (Figure 2c) (Mayasari et al., 2012). After the selection of G418-resistant ES cell colonies, we found that the number of colonies transduced with the sDT or dDT vector was 20- to 50-fold smaller than that with the UPATrap vector lacking DT (Figure 2c). The weaker toxin vectors of sDT176 and dDT176 produced a larger number of ES cell colonies than the wild-type DT counterpart vectors (Figure 2c), indicating that the weaker toxin, DT176 helped to trap genes and select clones, while wild-type DT was too toxic for ES cells to select clones.

Since the number of ES cell clones produced using DT vectors was very small, we selected the sDT176 and dDT176 vectors and performed further gene-trapping experiments. We obtained 357, 682, and 396 ES cell clones for trapping using the sDT176, dDT176, and non-DT UPATrap vectors, respectively, and found that the coding region had been trapped in 59.1, 65.1, and 72.9% of these clones with the sDT176, dDT176, and non-DT vectors, respectively (Figure 2d). The list of genes trapped by the DT, DT176, and UPATrap vectors was shown in Tables S2–S6. The other clones showed the integration of vectors at non-coding regions. Furthermore, the frequencies of integration into the sense (forward) strand at the coding region were higher in the dDT176 and non-DT vectors, while it was smaller for the sDT176 vector (Figure 2e), presumably due to the inappropriate selection of ES cells, the coding region of which was reversely trapped by the sDT176 vector, which cannot activate the expression of toxin proteins in inverse integration (Figure S1). While a removable exon trap (RET) vector showed a strong integration site bias toward the last intron of genes (Shigeoka et al., 2005), we found that UPATrap-derivative DT176 vectors showed similar integration in the regions of trapped genes between the first intron, gene body, and last intron (Figure 2f). This result provides support for the second portion of the gene-terminator cassette functioning in the termination of pre-mRNA transcription to suppress the inappropriate activation of the NMD pathway, resulting in similar integration within a gene.

3.3 | The double weaker toxin DT176 vector dominantly trapped silent genes in ES cells

To clarify whether the DT176 vector trapped transcriptionally silent genes in undifferentiated ES cells, we assessed the expression levels of trapped genes using the NCBI data sets of mRNA-driven EST libraries isolated from multiple murine ES cell datasets and an RNA-sequencing dataset from v6.5 murine ES cells (Guttman et al., 2010). We defined a transcriptionally silent gene by the absence of the corresponding EST and having smaller than 3 Reads Per Kilobase of exon per million (RPKM), which were defined in these murine ES cells, and found that the dDT176 vector trapped more genes showing silent and/or weak expression in ES cells than the non-DT UPATrap vector (69.6% versus 37.9% of trapped genes integrated in their forward orientation) (Figure 3a,b). In addition, dDT176-trapped ES cells showed significantly smaller RPKM of trapped genes than UPATrap-trapped ES cells (4.482 versus 11.040, p=0.0006) (Figure 3c). We then confirmed the expression levels of these silent genes in dDT176- and UPATrap-trapped ES cells (e.g., v6.4 and KY1.1 cell lines) by performing qRT-PCR on the representative genes (34 genes in UPATrap and 44 genes in DT176). In the silent genes examined, a correlation was observed between RPKM and qRT-PCR values in both vector-trapped cells (Figure 3d). dDT176-trapped cells showed significantly smaller qRT-PCR values for these genes than UPATrap-trapped cells (Figure 3e), indicating that the dDT176 vector dominantly trapped silent genes in ES cells. While five dDT176 trapped clones showed higher RPKM of the gene in gene expression dataset analysis (Figure 3c), among these clones, we confirmed the intra-vector deletions around the first DT cassette in 3 out of 5 clones by genomic PCR (Table S7). To elucidate the biological functions of trapped genes in ES cells, we performed a gene ontology (GO) analysis and found that dDT176 vector-trapped cells showed the enrichment of genes regulating non-ES and tissue-specific functions (e.g., neurons, blood cells, and muscle), which were not positively enriched in non-DT vector-trapped cells (Figure 3f), thereby supporting the ability of the DT176 vector to disrupt silent genes in ES cells, which are activated in the differentiation of ES cells and development.

3.4 | The DT176 vector deleted a gene in the differentiated cell lineage and labeled mutated cells

Due to the efficient trapping of the DT176 vector at silent genes in ES cells, we examined intra-vector integrity in genomes by performing genomic PCR on five regions of the vector prior to the induction of FLPo recombination (Figure 4a). The results obtained showed that 21 out of 24 dDT176-trapped clones exhibited intra-vector integrity, while 7 out of 8 dDT-trapped clones markedly lost most of the vector regions in cells, with only one out of 8 dDT-trapped clones showing integrity (Figure 4b and Table S7). Based on the high stability of the integrated DT176-vector region in these clones, we performed FLPo-mediated recombination on 13 DT-trapped ES clones, and found that they exhibited higher efficacies of recombination and the subsequent deletion of the vector region (average 73.2% from 22 to 100%) (Table S7). In order to perform targeted cell lineage gene ablation and labeling, we selected two DTrap vector-trapped clones (e.g., TM26-004 and TM28-026) and assessed the cellular functions of differentiated cells after gene deletion. The TM26-004 clone showed the integration of the DT vector in the gamma-crystallin E (Cryge) gene located at Chromosome 1 harboring the gamma-crystallin gene cluster region (Figure 4c), while the TM 28-026 clone showed the integration of the DT176 vector in the Perilipin-1 gene. Gamma crystallins are predominant proteins in the eye lens (Vendra, Khan, Chandani, Muniyandi, & Balasubramanian, 2016; Wistow, 2012), and genetic mutations in CRYG have been shown to induce cataracts in patients (Klopp, Loster, & Graw, 2001; Nag et al., 2007). We generated chimeric mice using the TM26-004-driven FLPo-treated clone, and bred mice to produce F1 mice showing EGFP-labeled lens (Figure 4e). In these mice, we did not find an impaired phenotype in lens that lacked the Cryg gene

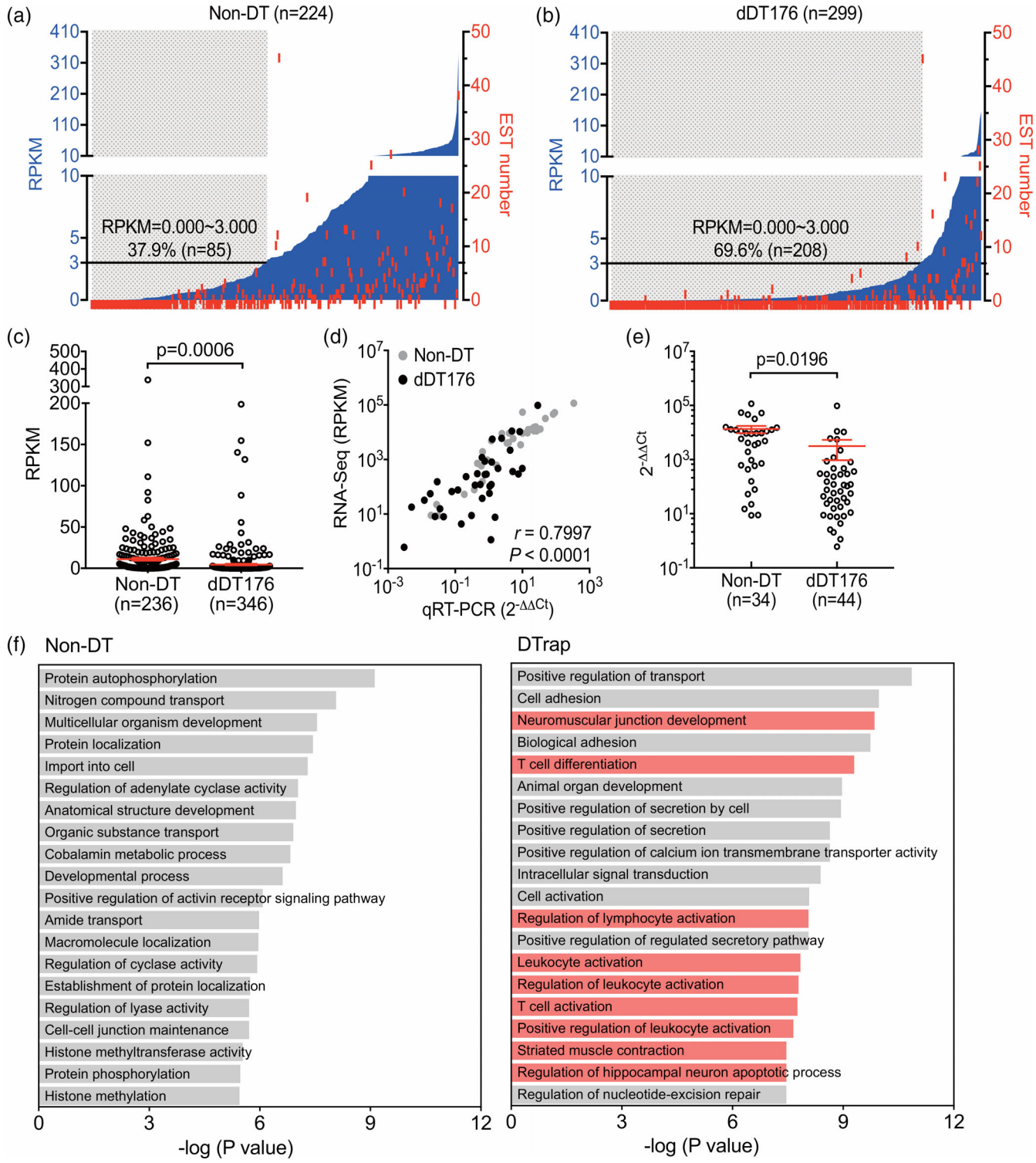

FIGURE 3 The double weaker toxin DT176 vector dominantly trapped silent genes in ES cells. (a,b) Expression levels of trapped genes defined by NCBI EST datasets and RNA-sequencing data in v6.5 ES cells showing silent genes occupied in 37.9% of trapped genes in the non-DT vector (a) and in 69.6% in the dDT176 vector (b). (c) Expression levels of vector-trapped genes defined by the RNA-sequencing dataset in v6.5 ES cells. Bars show the mean ± *SEM* and *p*-values analyzed by the Student's *t* test. (d) Correlation in the expression levels of selected genes trapped by the vectors between those in the RNA-sequencing dataset and those examined by quantitative RT-PCR in ES cells. (e) Expression levels of selected silent genes, which were defined by the published datasets, examined by quantitative RT-PCR in dDT176 vector- and non-DT vector-integrated ES cells. Bars show the mean ± *SEM* and *p*-values analyzed by the Student's *t* test. (f) Gene ontology (GO) analysis for trapped genes by the DT176 vector or non-DT vector

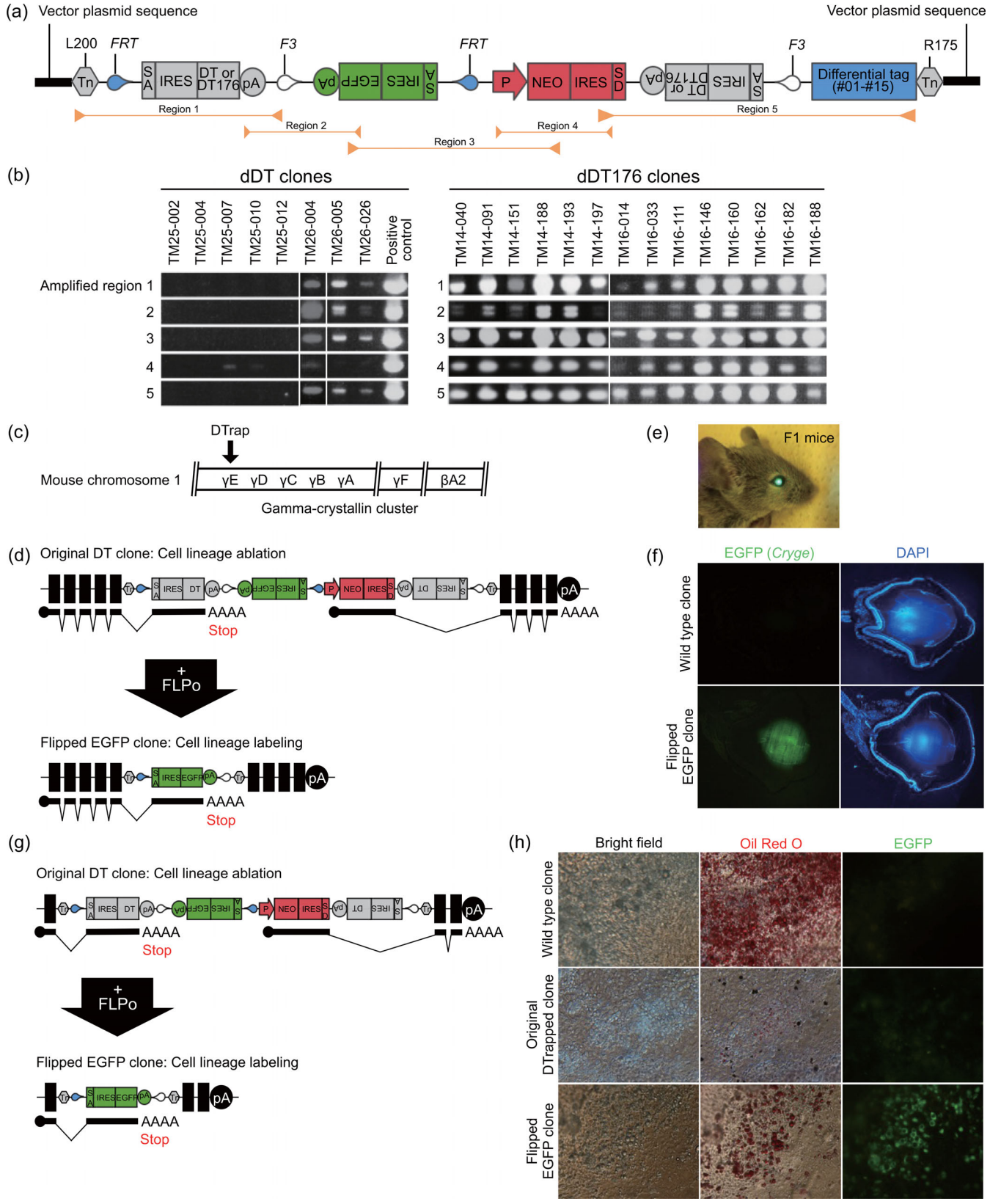

FIGURE 4 The DT176 vector deleted a gene in the differentiated cell lineage and labeled mutated cells. (a) Regions for evaluating integrated vector integrity in cells examined by genomic PCR. (b) Amplified PCR bands isolated from ES cell clones trapped with vectors showing the efficient and stable integration of the dDT176 vector. (c) Diagram of the Gamma-crystallin cluster in mouse chromosome 1. (d) Diagram of dDT vector integration in and FLPo recombination of Gamma-crystallin E (Cryge) in the TM26-004 clone. Tag sequence boxes were omitted in (d,g). (e) Representative picture of a F1 mouse generated from the TM26-004-driven FLPo-treated clone showing EGFP-expressing green eye. (f) Lens sections of a wild-type mouse and F1 mouse from the TM26-004 clone with FLPo-induced recombination. (g) Diagram of dDT176 vector integration in and FLPo recombination of Perilipin-1 in the TM28-026 clone. (h) Representative pictures of adipocytes examined by Oil red O staining induced from the TM28-026 clone with or without FLPo-induced recombination and control ES cells upon in vitro adipocyte differentiation

accompanied with the expression of EGFP (Figure 4f), presumably due to the compensatory function of the other crystallin family genes in lens. Perilipins function to coat the surface of lipid droplets in adipocytes and regulate lipolysis (Sztalryd & Brasaemle, 2017). Perilipin knockout mice showed reduced amounts of adipose tissue and activated lipolysis in their adipocytes (Arimura, Horiba, Imagawa, Shimizu, & Sato, 2004). We induced the adipocyte differentiation of the TM28-026 and FLPo-induced EGFP clones under in vitro conditions (Figure 4g). The TM28-026 ES clone did not produce adipocytes due to the induction of the expression of toxins upon the activation of perilipin expression in differentiation (Figure 4h). FLPo-induced EGFP clones lacking the DT176 toxin showed the weaker differentiation of adipocytes, as examined by Oil Red staining, accompanied by EGFP expression than control cells, which showed the abundant accumulation of adipocytes under this experimental setting (Figure 4h), indicating the successful deletion of the perilipin gene in the adipocyte lineage labeled by EGFP.

4 | DISCUSSION

Although advances have been achieved in mutagenesis-based gene trapping methodologies, many transcriptionally silent genes have yet to be targeted because of the difficulties associated with targeting silent genes and selecting properly trapped clones (Mayasari et al., 2012). In the last 5 years, the CRISPR-Cas9 system has been utilized for mutagenesis and has become more popular than conventional gene trapping because of the convenience of deleting and/or editing genomic regions including genes and the higher efficiency of targeting (Horvath & Barrangou, 2010; Terns & Terns, 2011; Wiedenheft, Sternberg, & Doudna, 2012). However, CRISPR-Cas9 frequently induces off-target mutagenesis regardless of the distance from the target region (Fu et al., 2013), and does not effectively detect off-target mutations due to the editing of CRISPR-Cas9 unless the whole-genome sequencing of a cell is performed. In the present study, we developed a novel *Tol*2 transposon-based DTrap vector that dominantly integrated into a single genetic region in the genome, while maintaining the high integrity of the integrated region, but also efficiently selected the silent gene-trapped clone due to weaker and double-oriented DT cassettes. In addition, DTrap vector integrated into the unidentified non-coding regions with relatively high frequency, which may include long noncoding RNA. Since we also successfully traced the cell lineage upon the deletion of genes in development for eye lens, adipocytes and other lineages (data not shown) under in vitro and in vivo settings, the dDT176 vector provides a pure forward genetic approach in mammalian stem cells, which will provide a more detailed understanding of the physiological roles of coding- and non-coding genetic regions under in vitro and in vivo conditions in a high throughput and genome-wide screening manner.

Induced pluripotent stem (iPS) cells have been utilized to create human cancer models that provide opportunities for basic and translational cancer research (Papapetrou, 2016). Since patient-derived iPS cells capture genetic and epigenetic alterations in cancer cells (Ohnishi et al., 2014), DTrap may be applied to the identification of cooperative genetic regions/mutations via genome-wide molecular analyses and genotype–phenotype studies upon the differentiation of iPS cells into cancer cells harboring the dDT176 integration that will be monitored by the presence of EGFP. The DTrap vector may also empower translational research through identification of therapeutic targets and biomarkers. Furthermore, the DTrap vector may empower future translational research, such as the identification of therapeutic targets and biomarkers in cancer cells; however, further studies are needed to assess the feasibility of DTrap in cancer models for translational cancer research under in vivo settings.

In this study, we demonstrate that dDTrap provides a systematic forward genetic approach by targeting transcriptionally silent genes in ES cells, which will provide novel insights into the molecular mechanisms of genes that play crucial roles in development under in vitro and in vivo settings. Based on that the DT176 vector deleted a gene and successfully labeled mutated cells upon the differentiation such as adipocytes, we believe that the dDTrap vector will improve feasibility of developmental genetics via performing a precisely targeting and high throughput screening of genes, of which differential expression create different cell types from identical ES cells.

ACKNOWLEDGMENTS

We thank Koichi Kawakami (National Institute of Genetics, Japan), Kenji Kohno (Nara Institute of Science and Technology, Japan), A. Francis Stewart (TU Dresden, Germany), and Junji Takeda (Osaka University, Osaka, Japan) for kindly providing us with the *Tol2*-associated plasmids, pCRM176 plasmid, pCAGGS-FLPo-IRES-Puro^r-poly (A), and KY1.1 ES cells, respectively.

CONFLICT OF INTEREST

The authors declare that there are no potential conflicts of interest to disclose.

ORCID

Jie Bai https://orcid.org/0000-0002-7485-4450

REFERENCES

Arimura, N., Horiba, T., Imagawa, M., Shimizu, M., & Sato, R. (2004). The peroxisome proliferator-activated receptor gamma regulates expression of the perilipin gene in adipocytes. *The Journal of Biological Chemistry*, *279*(11), 10070–10076.

Austin, C. P., Battey, J. F., Bradley, A., Bucan, M., Capecchi, M., Collins, F. S., … Zambrowicz, B. (2004). The knockout mouse project. *Nature Genetics*, *36*(9), 921–924.

Carver, T., Harris, S. R., Berriman, M., Parkhill, J., & McQuillan, J. A. (2012). Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics*, *28*(4), 464–469.

Dani, C., Smith, A. G., Dessolin, S., Leroy, P., Staccini, L., Villageois, P., … Ailhaud, G. (1997). Differentiation of embryonic stem cells into adipocytes in vitro. *Journal of Cell Science*, *110*(11), 1279–1285.

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K., & Sander, J. D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature Biotechnology*, *31*(9), 822–826.

Gossler, A., Joyner, A. L., Rossant, J., & Skarnes, W. C. (1989). Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science*, *244*(4903), 463–465.

Guttman, M., Garber, M., Levin, J. Z., Donaghey, J., Robinson, J., Adiconis, X., … Regev, A. (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nature Biotechnology*, *28*(5), 503–510.

Horvath, P., & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science*, *327*(5962), 167–170.

Ishida, Y., & Leder, P. (1999). RET: A poly A-trap retrovirus vector for reversible disruption and expression monitoring of genes in living cells. *Nucleic Acids Research*, *27*(24), e35–e335.

Kawakami, K., Shima, A., & Kawakami, N. (2000). Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(21), 11403–11408.

Klopp, N., Loster, J., & Graw, J. (2001). Characterization of a 1-bp deletion in the gammaE-crystallin gene leading to a nuclear and zonular cataract in the mouse. *Investigative Ophthalmology and Visual Science*, *42*(1), 183–187.

Koga, A., Suzuki, M., Inagaki, H., Bessho, Y., & Hori, H. (1996). Transposable element in fish. *Nature*, *383*(6595), 30.

Kohno, K., & Uchida, T. (1987). Highly frequent single amino acid substitution in mammalian elongation factor 2 (EF-2) results in expression of resistance to EF-2-ADP-ribosylating toxins. *The Journal of Biological Chemistry*, *262*(25), 12298–12305.

Kohno, K., Uchida, T., Ohkubo, H., Nakanishi, S., Nakanishi, T., Fukui, T., … Okada, Y. (1986). Amino acid sequence of mammalian elongation factor 2 deduced from the cDNA sequence: Homology with GTP-binding proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(14), 4978–4982.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Matsuda, E., Shigeoka, T., Iida, R., Yamanaka, S., Kawaichi, M., & Ishida, Y. (2004). Expression profiling with arrays of randomly disrupted genes in mouse embryonic stem cells leads to in vivo functional analysis. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(12), 4170–4174.

Maxwell, F., Maxwell, I. H., & Glode, L. M. (1987). Cloning, sequence determination, and expression in transfected cells of the coding sequence for the tox 176 attenuated diphtheria toxin A chain. *Molecular and Cellular Biology*, *7*(4), 1576–1579.

Mayasari, N. I., Mukougawa, K., Shigeoka, T., Kawakami, K., Kawaichi, M., & Ishida, Y. (2012). Mixture of differentially tagged Tol2 transposons accelerates conditional disruption of a broad spectrum of genes in mouse embryonic stem cells. *Nucleic Acids Research*, *40*(13), e97.

McMahon, A. P., & Bradley, A. (1990). The Wnt-1 (int-1) proto-oncogene is required for development of a large region of the mouse brain. *Cell*, *62*(6), 1073–1085.

Nag, N., Peterson, K., Wyatt, K., Hess, S., Ray, S., Favor, J., … Wistow, G. (2007). Endogenous retroviral insertion in Cryge in the mouse No3 cataract mutant. *Genomics*, *89*(4), 512–520.

Niwa, H., Araki, K., Kimura, S., Taniguchi, S., Wakasugi, S., & Yamamura, K. (1993). An efficient gene-trap method using poly A trap vectors and characterization of gene-trap events. *Journal of Biochemistry*, *113*(3), 343–349.

Ohnishi, K., Semi, K., Yamamoto, T., Shimizu, M., Tanaka, A., Mitsunaga, K., … Yamada, Y. (2014). Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. *Cell*, *156*(4), 663–677.

Papapetrou, E. P. (2016). Patient-derived induced pluripotent stem cells in cancer research and precision oncology. *Nature Medicine*, *22*(12), 1392–1401.

Rubin, C. S., Hirsch, A., Fung, C., & Rosen, O. M. (1978). Development of hormone receptors and hormonal responsiveness in vitro. Insulin receptors and insulin sensitivity in the preadipocyte and adipocyte forms of 3T3-L1 cells. *The Journal of Biological Chemistry*, *253*(20), 7570–7578.

Scherdin, U., Rhodes, K., & Breindl, M. (1990). Transcriptionally active genome regions are preferred targets for retrovirus integration. *Journal of Virology*, *64*(2), 907–912.

Shigeoka, T., Kawaichi, M., & Ishida, Y. (2005). Suppression of nonsense-mediated mRNA decay permits unbiased gene trapping in mouse embryonic stem cells. *Nucleic Acids Research*, *33*(2), e20.

Stanford, W. L., Cohn, J. B., & Cordes, S. P. (2001). Gene-trap mutagenesis: Past, present and beyond. *Nature Reviews. Genetics*, *2*(10), 756–768.

Sztalryd, C., & Brasaemle, D. L. (2017). The perilipin family of lipid droplet proteins: Gatekeepers of intracellular lipolysis. *Biochimica et Biophysica Acta – Molecular and Cell Biology of Lipids*, *1862*(10 Pt B), 1221–1232.

Terns, M. P., & Terns, R. M. (2011). CRISPR-based adaptive immune systems. *Current Opinion in Microbiology*, *14*(3), 321–327.

Uchida, T., Pappenheimer, A. M., Jr., & Greany, R. (1973). Diphtheria toxin and related proteins. I. Isolation and properties of mutant proteins serologically related to diphtheria toxin. *The Journal of Biological Chemistry*, *248*(11), 3838–3844.

Urasaki, A., Morvan, G., & Kawakami, K. (2006). Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition. *Genetics*, *174*(2), 639–649.

Vendra, V. P., Khan, I., Chandani, S., Muniyandi, A., & Balasubramanian, D. (2016). Gamma crystallins of the human eye lens. *Biochimica et Biophysica Acta*, *1860*(1 Pt B), 333–343.

Wiedenheft, B., Sternberg, S. H., & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, *482*(7385), 331–338.

Wistow, G. (2012). The human crystallin gene families. *Human Genomics*, *6*, 26.

Wu, X., Li, Y., Crise, B., & Burgess, S. M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science*, *300*(5626), 1749–1751.

Yamaizumi, M., Mekada, E., Uchida, T., & Okada, Y. (1978). One molecule of diphtheria toxin fragment A introduced into a cell can kill the cell. *Cell*, *15*(1), 245–250.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Bai J, Kondo R, Mayasari NI, et al. Diphtheria toxin-mediated transposon-driven poly (A)-trapping efficiently disrupts transcriptionally silent genes in embryonic stem cells. *genesis*. 2020;58:e23386. https://doi.org/10.1002/dvg.23386